(4)

FINAL REPORT

ONR Contract Number N00014-87-K-0184
Richard M. Fujimoto
Computer Science Department
University of Utah

1 Contract Information

Contract Title: Performance Directed Simulation

Contract Amount: $100,000

Contract Duration: February 15, 1987 through December 31, 1988

Principal Investigator: Richard M. Fujimoto

Electronic Mail Address: fujimoto@cs.utah.edu

Institution: University of Utah

Scientific Officer: Andre van Tilborg

2 Scientific Goals and Results

The overall scientific objective of this research was to develop techniques for speeding up discrete event simulation (DES) programs in applications such as simulation of computer architectures, communication networks, battlefield scenarios, etc. Such simulations often require prohibitive amounts of time on conventional computing hardware.

Several major successes and important new results were obtained during the grant period in this problem domain. The most significant results are summarized ~~below.~~ in the following areas.

2.1 Optimistic, Conservative Synchronization Mechanisms → See 3

Parallel simulation techniques broadly fall into two categories: optimistic and conservative mechanisms. Substantial new results were developed in each of these areas. The most significant results regarding optimistic methods (Time Warp) were [T1]:

- A new technique called *direct cancellation* was developed that substantially reduces many of the overheads in simulations using Time Warp. This mechanism was implemented in a version of Time Warp that is now operational on a BBN Butterfly multiprocessor.
- Using direct cancellation, substantial speedups of several benchmark programs were obtained. In particular, speedups as high as 56.8 using 64 processors were observed for a simulation of a queuing network configured as an eight dimensional hypercube. To our knowledge, this is the best speedup that has ever been recorded for a parallel discrete event simulation program.
- Time Warp was observed to achieve substantial speedups for applications where conservative mechanisms based on deadlock avoidance and deadlock detection and recovery have substantial difficulty (also, see below).

- Several qualitative results were discovered regarding the behavior of the Time Warp mechanism. In particular, the ability to rapidly cancel erroneous computations is crucial to achieving good performance. This is the essential reason that direct cancellation yields such good performance.

- Although good performance has been obtained for the benchmarks that we have examined thus far, other experiments indicate that state saving overheads lead to serious degradations in performance, even for programs containing a modest amount of state. These results highlight the need for hardware support to assist in the state management problem. Such hardware support has been studied extensively, as discussed below.

2.2 Conservative Synchronization Mechanisms

Both synthetic and actual workloads have been used to experiment with conservative simulation mechanisms using deadlock avoidance, and deadlock detection and recovery. The principal results of this work include [T11, T13]:

- Several techniques were developed and implemented to use shared memory to optimize performance of these simulation algorithms.

- The lookahead ability of logical processes plays a critical role in determining the efficiency of these simulation algorithms. Lookahead refers to the extent to which the simulation can predict what will happen in the future based on what has already happened in the past. Good performance can only be obtained if there is a high degree of parallelism, and the application has good lookahead properties. This tends to make performance highly sensitive to minor variations in the system being simulated. Performance using Time Warp appears not to have this problem.

- A phenomenon that we call *message avalanche* was observed in the deadlock detection and recovery simulator for moderate to high degrees of parallelism, and was necessary to achieve efficient execution. The poorer the lookahead ability of a process, the greater the amount of parallelism necessary to achieve avalanche. If lookahead is sufficiently poor, avalanche may never be observed for workloads of practical interest.

- Deadlock detection and recovery simulators containing different types of logical processes can be adversely affected by a small number of processes that exhibit poor lookahead ability. The existence of a few such processes can greatly increase the amount of parallelism necessary to achieve avalanche, even if many other processes contain very good lookahead properties. The deadlock avoidance simulator is not as severely affected by this behavior if the bulk of the simulation activity avoids processes with poor lookahead.

- Queuing networks that contain cycles, previously thought to be ill-suited for conservative distributed simulation algorithms, can achieve good performance if servers are reprogrammed to take advantage of all available lookahead. In some cases, as much as an order in magnitude performance improvement was observed by simply recoding to fully exploit lookahead.

- Simulation applications such as those containing infrequent preemptive events *inherently* have poor lookahead properties, and appear ill-suited for the deadlock avoidance and deadlock recovery algorithms. Applications containing state dependent behavior (e.g., load balancing mechanisms) similarly contain moderate to poor lookahead properties.

- Simulations of several hypercube-based communication networks with varying degrees of lookahead provide empirical data to support the above conclusions. In particular, simulators with progressively poorer lookahead properties (the first used FCFS queues, the second prioritized messages but with no preemption, and the third used preemption) were observed to achieve progressively poorer performances.

- Moderate to high amounts of parallelism were required to obtain good performance. Not unexpectedly, the amount of parallelism required for efficient execution increases with node degree.

Although experimentation is not yet complete, our initial results point to Time Warp as a far more promising approach to parallel discrete event simulation than existing conservative techniques.

2.3 Special Purpose Hardware for Discrete Event Simulation

"Optimistic" parallel DES algorithms such as Time Warp offer an attractive alternative to "conservative" strategies. One problem that threatens the viability of rollback-based methods is the necessity of frequent state saving. This represents a substantial overhead, particularly for programs containing large amounts of state (for example, embedded continuous simulations). Work in this area has been focused on the *rollback chip*, a component to efficiently implement state saving and rollback functions in hardware. Principal results of this work include [T2, T3, T4, T5, T6, T7, T15]:

- Highly efficient mechanisms for state saving and rollback have been defined, and simulated to verify functional correctness.

- Performance results (obtained from the simulator) indicate that the cost of state saving using the rollback chip is projected to be only a few percent of processor performance, even when the size of state is large (e.g., several megabytes) and state save operations occur frequently (e.g., every 100 microseconds).

- Several portions of the rollback chip have been designed. A design using off the shelf components is nearing completion. Layout of custom chips for a special component used by the rollback chip's cache is also nearly complete.

- Recently, work has been focused on a special purpose discrete event simulation engine based on a shared-memory architecture. Based on the rollback chip work, a key feature of this processor is a specialized memory system that is addressed by *time* (simulated time that is) as well as space, allowing processors to examine *past* values of variables. This memory system provides direct support for interactions among processes that, at any moment of the computation, have advanced to different points in simulated time. This is the normal mode of operation in parallel discrete event simulation.

- At the time this grant terminated, work was focused on a new type of *general-purpose* parallel computer architecture called the *Virtual Time Machine (VTM)* [T2]. The VTM architecture is based on our claim that it is difficult to effectively utilize existing parallel computers because today's machine architectures do not adequately represent *time*. Specifically, VTM uses the aforementioned space-time memory system to detect violations of data dependence relationships at *runtime*, and *automatically* recovers from them. We believe this new approach will greatly facilitate the automatic parallelization of sequential programs, and offers great potential for addressing many difficult, long-standing problems in parallel computation today.

2.4 Other Results

Other work that has been partially or completely funded under this grant include:

- Work has continued in developing efficient, shared-memory based synchronization algorithms for rendezvous constructs (e.g., the generalized alternative construct of CSP) [T9, T12]. A proof of the algorithm was completed last year. This year, we have focused attention on a prototype implementation that was developed for the Butterfly. Now complete, the prototype was instrumented and refined to optimize performance.

- Techniques using *direct execution* (as opposed to interpretation) have been developed to provide efficient, instruction level simulation models for CPUs [T14]. Such techniques can be used to achieve efficient simulations of single and multiple processor computer systems, providing one or more orders of magnitude speed up over conventional techniques. Results for a prototype implementation for simulation models for 68000 based processors were reported in the current year.

- Work has also been performed in cooperation with Dr. Ganesh Gopalakrishnan (who is supported by a grant from NSF) in obtaining practical usage of the hardware description language HOP which Dr. Gopalakrishnan has developed [T8, T10]. Portions of the rollback chip were specified in HOP. Emphasis in this work has been towards formal verification of hardware systems.

- A method for determining the optimal performance of a discrete event simulation program was developed. This program uses an oracle to guide the simulation, and is useful for measuring the amount of parallelism available in a simulation application [T16].

3 Participants

The principal investigator is Dr. Richard M. Fujimoto (University of Utah). Much of the hardware design work has been done in collaboration with Dr. Ganesh Gopalakrishnan who is also on the faculty at Utah.

The following students have been involved, to some extent, in the project (total: 4 PhD, 7 MS, 1 undergrad): Venkatesh Akella (PhD), Hwa-Chung Feng (MS), William Campbell (MS), Armin Liebchen (undergrad), Narayana Mani (MS), Surya Mantha (PhD), Gautam Mehrotra (MS), Kevin Smith (MS), Kay Smith (MS), Steve Swope (PhD), Rich Thomson (PhD), and Jya-Jang Tsai (MS). Among these, Feng, Swope, Thomson, and Tsai received partial support through research assistantships.

During the contract period, two MS degrees and one BS degree were completed:

Hwa-Chung Feng, M.S. in Computer Science, "A Multiprocessor Implementation of CSP," December 1987.

Jya-Jang Tsai, M.S. in Computer Science, "The Design and Performance of the Rollback Chip: Hardware Support for Time Warp," September 1988.

Armin Liebchen, B.S. in Computer Science, June 1988.

4 List of Publications, Reports, and Presentations

4.1 Papers Published in Refereed Journals

[J1] R. M. Fujimoto and W. B. Campbell, "Efficient Instruction Level Simulation of Computers," *Transactions of the Society for Computer Simulation*, vol. 5, No. 2, April 1988.

[J2] R. M. Fujimoto and H. C. Feng, "A Shared Memory Algorithm and Proof for the Generalized Alternative Construct in CSP," *International Journal of Parallel Processing*, vol. 16, No. 3, June 1987.

[J3] R. M. Fujimoto, "Performance Measurements of Distributed Simulation Strategies," *Transactions of the Society for Computer Simulation*, to appear (also University of Utah Technical Report UUCS-87-026a).

4.2 Papers Submitted to Refereed Journals

The following have been submitted, and are currently under review:

[J4] R. M. Fujimoto and P. J. Waterman, "The Importance of Lookahead in Conservative Parallel Discrete Event Simulation," submitted to *Simulation*.

[J5] R. M. Fujimoto, J. Tsai and G. Gopalakrishnan, "Design and Evaluation of the Rollback Chip: Special Purpose Hardware for Time Warp," submitted to *IEEE Transactions on Computers* (also University of Utah Technical Report UUCS-88-011).

[J6] G. Gopalakrishnan, V. Akella, and R. M. Fujimoto, "HOP: A Language, Process Model, and Design System for Hardware: Process Composition Techniques and Applications for VLSI Design," submitted to *Integration*.

[J7] R. M. Fujimoto, "Time Warp on a Shared Memory Multiprocessor," submitted to *Transactions of the Society for Computer Simulation*, (also University of Utah Technical Report UUCS-88-021a).

4.3 Books and Chapters in Books

[B1] D. A. Reed and R. M. Fujimoto, *Multicomputer Networks: Message-Based Parallel Processing*, MIT Press, Nov. 1987 (book).

[B2] G. Gopalakrishnan, R. M. Fujimoto, V. Akella, N. Mani, and K. Smith, "Specification Driven Design of Custom VLSI Architectures," in Specification Driven VLSI Design, Springer-Verlag, 1988 (chapter in book).

[B3] B. Unger and R. M. Fujimoto (editors), "Distributed Simulation," Society for Computer Simulation, to appear, 1989 (book).

4.4 Refereed Conference Papers

[C1] R. M. Fujimoto, "Lookahead in Parallel Discrete Event Simulation," 1988 International Conference on Parallel Processing, August 1988.

[C2] H. C. Feng and R. M. Fujimoto, "A Shared Memory Algorithm and Performance Evaluation of the Generalized Alternative Construct," 1988 International Conference on Parallel Processing, August 1988.

[C3] R. M. Fujimoto, J. Tsai and G. Gopalakrishnan, "Design and Performance of Special Purpose Hardware for Time Warp," 15th Annual International Symposium on Computer Architecture, June 1988.

[C4] R. M. Fujimoto, "Performance Measurements of Distributed Simulation Programs," 1988 SCS Multiconference: Distributed Simulation, February 1988 (also University of Utah Technical Report UUCS-87-026).

[C5] R. M. Fujimoto, J. J. Tsai and G. Gopalakrishnan, "The Roll Back Chip: Hardware Support For Distributed Simulation Using Time Warp," 1988 SCS Multiconference: Distributed Simulation, February 1988 (also University of Utah Technical Report UUCS-87-025).

[C6] R. M. Fujimoto and W. B. Campbell, "Direct Execution Models of Processor Behavior and Performance," 1987 Winter Simulation Conference, December 1987.

[C7] S. Swope and R. M. Fujimoto, "Optimal Performance of Distributed Simulation Programs," 1987 Winter Simulation Conference, December 1987 (also University of Utah Technical Report UUCS-87-023).

4.5 Papers Submitted to Refereed Conferences

[C8] R. M. Fujimoto, "Time Warp on a Shared Memory Multiprocessor," submitted to *1989 International Conference on Parallel Processing* (also University of Utah Technical Report UUCS-88-021a).

[C9] R. M. Fujimoto, "The Virtual Time Machine," submitted to *1989 ACM Symposium on Parallel Algorithms and Architectures*, (also University of Utah Technical Report UUCS-88-019).

4.6 Technical Reports

[T1] R. M. Fujimoto, "Time Warp on a Shared Memory Multiprocessor," University of Utah Technical Report UUCS-88-021a, January 1989 (submitted to *Transactions of the Society for Computer Simulation*).

[T2] R. M. Fujimoto, "The Virtual Time Machine," University of Utah Technical Report UUCS-88-019, January 1989 (submitted to *ACM Symposium on Parallel Computer Algorithms and Architectures*).

[T3] J. Tsai, "The Design and Performance of the Rollback Chip: Hardware Support for Time Warp", M.S. Thesis, University of Utah Technical Report UUCS-88-021, September 1988.

[T4] R. Thomson "Design Specifications for Hardware Assisted Rollback Computation," University of Utah Technical Report UUCS-88-014, September 1988.

[T5] V. Akella, "YAMA: Yet Another Microassembler Description and User's Guide," University of Utah Technical Report UUCS-88-016, September 1988.

[T6] N. Mani and S. Mantha, "The Design and Implementation of a Special Purpose Cache for the Rollback Chip," University of Utah Technical Report UUCS-88-015, September 1988.

[T7] R. M. Fujimoto, J. Tsai and G. Gopalakrishnan, "Design and Evaluation of the Rollback Chip: Special Purpose Hardware for Time Warp," University of Utah Technical Report UUCS-88-011, July 1988 (submitted to *IEEE Transactions on Computers*).

[T8] G. Gopalakrishnan, R. M. Fujimoto, V. Akella, and N. Mani, "HOP: A Process Model for Synchronous Hardware: Semantics and Experiments in Process Composition," University of Utah Technical Report UUCS-88-012, August 1988.

[T9] H. C. Feng, "A Multiprocessor Implementation of CSP," M.S. Thesis, University of Utah Technical Report UUCS-88-022, March 1988.

[T10] G. Gopalakrishnan and R. M. Fujimoto, "HOP: A Process Model for Synchronous Hardware Systems," University of Utah Technical Report UUCS-88-003, Feb. 1988.

[T11] R. M. Fujimoto, "Performance Measurements of Distributed Simulation Strategies," University of Utah Technical Report UUCS-87-026a, Nov. 1987 (to appear in *Transactions of the Society for Computer Simulation*).

[T12] R. M. Fujimoto and H. C. Feng, "A Shared Memory Algorithm and Proof for the Generalized Alternative Construct in CSP," University of Utah Technical Report UUCS-87-022, Sept. 1987 (appeared in *International Journal of Parallel Processing*, vol. 16, No. 3, June 1987).

[T13] R. M. Fujimoto, "Performance Measurements of Distributed Simulation Programs," University of Utah Technical Report UUCS-87-026, Sept. 1987 (appeared in 1988 SCS Multiconference: Distributed Simulation, February 1988).

[T14] R. M. Fujimoto and W. B. Campbell, "Efficient Instruction Level Simulation of Computers," University of Utah Technical Report UUCS-87-021, Sept. 1987 (appeared in *Transactions of the Society for Computer Simulation*, vol. 5, No. 2, April 1988).

[T15] R. M. Fujimoto, J. J. Tsai and G. Gopalakrishnan, "The Roll Back Chip: Hardware Support For Distributed Simulation Using Time Warp," University of Utah Technical Report UUCS-87-025, Sept. 1987 (appeared in 1988 SCS Multiconference: Distributed Simulation, February 1988).

[T16] S. Swope and R. M. Fujimoto, "Optimal Performance of Distributed Simulation Programs," University of Utah Technical Report UUCS-87-023, Sept. 1987 (appeared in 1987 Winter Simulation Conference, December 1987).

4.7 Invited Presentations

1. R. M. Fujimoto, "Design and Performance of the Rollback Chip," University of Calgary, July 1988.
2. R. M. Fujimoto, "Specification Driven Design of Custom VLSI Architectures," Workshop on Hardware Verification, Banff, Canada, July 1988.
3. R. M. Fujimoto, "Design and Performance of the Rollback Chip," University of California at Los Angeles, May 1988.

4. R. M. Fujimoto, "Parallel Discrete Event Simulation," ONR Workshop on Parallel Computation, Washington D.C., May 1988.

5. R. M. Fujimoto, "Performance of Distributed Simulation Algorithms," University of Calgary, February, 1988.

6. R. M. Fujimoto, "Distributed Simulation Experiments on the BBN Butterfly Multiprocessor," BBN Users Group Meeting, Cambridge, Massachusetts, Nov. 1987.

4.8 Contributed Presentations

1. R. M. Fujimoto, "Lookahead in Parallel Discrete Event Simulation," 1988 International Conference on Parallel Processing, August 1988.

2. H. C. Feng, "A Shared Memory Algorithm and Performance Evaluation of the Generalized Alternative Construct," 1988 International Conference on Parallel Processing, August 1988.

3. R. M. Fujimoto, "Design and Performance of Special Purpose Hardware for Time Warp," 15th Annual International Symposium on Computer Architecture, June 1988.

4. R. M. Fujimoto, "Performance Measurements of Distributed Simulation Programs," 1988 SCS Multiconference: Distributed Simulation, February 1988.

5. R. M. Fujimoto, "The Roll Back Chip: Hardware Support For Distributed Simulation Using Time Warp," 1988 SCS Multiconference: Distributed Simulation, February 1988.

6. W. B. Campbell, "Direct Execution Models of Processor Behavior and Performance," 1987 Winter Simulation Conference, December 1987.

7. S. Swope, "Optimal Performance of Distributed Simulation Programs," 1987 Winter Simulation Conference, December 1987.

5 List of Awards

1. R. M. Fujimoto (University of Utah), "Outstanding Teaching Award," awarded by College of Engineering, University of Utah.

6 Publications, Patents, Honors (Numbers Only)

The following figures cover the entire 19 months of the contract period.

- Papers submitted to refereed journals (not yet published): 4
- Papers published in refereed journals: 3
- Papers submitted to refereed conferences: 2
- Papers published in refereed conferences: 7
- Books submitted for publication: 0
- Books published: 2 books (1 as author, 1 as editor), 1 chapter in book

- Patents filed: 0
- Patents granted: 0
- Invited presentations (including colloquium): 6
- Contributed presentations: 7
- Honors, awards, prizes: 1
- Technical reports published: 16
- Number of graduate students: total: 4 PhD, 7 MS; 4 supported as RAs; 2 MS students graduated.
- Number of post docs: 0